MARKED UP VERSION OF AMENDMENTS

Specification Amendments Under 37 C.F.R. § 1.121(b)(1)(iii)

Replace the paragraph at page 11, lines 1-3 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Figures 2A-B [is] are [graph of] scatterplots showing a neighborhood analysis of genes correlating to Acute Lymphoblastic Leukemia (ALL; Figure 2A) or Acute Myeloid Leukemia (AML; Figure 2B).

Replace the paragraph at page 11, lines 9-13 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Figures 4A-B [is] are a set of graphs showing neighborhood analysis of genes in AML samples from patients with different clinical responses to treatment. Results are shown for 15 AML samples for which long-term clinical follow-up was available, with genes more highly expressed in the treatment failure group in [the left panel] Figure 4A and genes more highly expressed in the treatment success group in [the right panel] Figure 4B.

Replace the paragraph at page 12, lines 10-11 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

Figures 11A-D [is] are [an] illustrations showing the assessment of statistical significance of gene-class correlations using neighborhood analysis.

Replace the paragraph at page 41, line 20 through page 42, line 9 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

The 38 acute leukemia samples were subjected to neighborhood analysis and revealed a strikingly high density of genes correlated with the AML-ALL distinction. Roughly 1100 genes were more highly correlated with the AML-ALL class distinction than would be expected by chance (Figs. 2A-B). Figures 2A-B show[s] the number of genes within various 'neighborhoods' of the ALL/AML class distinction together with curves showing the 5% and 1% significance levels for the number of genes within corresponding neighborhoods of the randomly permuted class distinctions. Genes more highly expressed in ALL compared to AML are shown in the left panel; those more highly expressed in AML compared to ALL are shown in right panel. Note the large number of genes highly correlated with the class distinction. In the left panel (higher in ALL), the number of genes with correlation $P(g,c) > 0.30$ was 709 for the AML-ALL distinction, but had a median of 173 genes for random class distinctions. Note that $P(g,c) = 0.30$ is the point where the observed data intersects the 1% significance level, meaning that 1% of random neighborhoods contain as many points as the observed neighborhood round the AML-ALL distinction. Similarly, in the right panel (higher in AML), 711 genes with $P(g,c) > 0.28$ were observed, whereas a median of 136 genes is expected for random class distinctions.

Replace the paragraph at page 46, lines 4-20 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

The choice to use 50 informative genes in the predictor was somewhat arbitrary, although well within the total number of genes strongly correlated with the class distinction (Figs. 2A-B). In fact, the results proved to be quite insensitive to this choice: class predictors based on between 10 and 200 genes were tested and all were found to be 100% accurate, reflecting the strong correlation of genes with the AML-ALL distinction Although the number of genes used had no significant effect on the outcome in this case (median PS for cross-validation ranged from 0.81 to 0.68 over a range of predictors employing 10-200 genes, all with 0% error), it may matter in other instances. One approach is to vary the number of genes used, select the number that maximizes the accuracy rate in cross-validation and then use the resulting model on the independent dataset. In any case, it is recommend that at at least 10 genes be used for two reasons. Class predictors employing a small number of genes may depend too heavily on any one

gene and can produce spuriously high prediction strengths (because a large 'margin of victory' can occur by chance due to statistical fluctuation resulting from a small number of genes). In general, the 1% confidence line in neighborhood analysis was also considered to be the upper bound for gene selection.

Replace the paragraph at page 46, line 20 through page 49, line 10 with the below paragraph marked up by way of bracketing and underlining to show the changes relative to the previous version of the paragraph.

The ability to predict response to chemotherapy among the 15 adult AML patients who had been treated with an anthracycline-cytarabine regimen and for whom long-term clinical follow-up was available was explored. Treatment failure was defined as failure to achieve a complete remission following a standard induction regimen including 3 days of anthracycline and 7 days of cytarabine. Treatment successes were defined as patients in continuous complete remission for a minimum of 3 years. FAB subclass M3 patients were excluded, but samples were otherwise not selected with regard to FAB criteria. Eight patients failed to achieve remission following induction chemotherapy, while the remaining seven patients remain in remission for 46-84 months. In contrast to the situation for the AML-ALL distinction, neighborhood analysis found no striking excess of genes correlated with response to chemotherapy (Figs 4A-B). The data fall close to the mean expected from random clusters. Nonetheless, the single most highly correlated gene, HOXA9 (arrow), is biologically related to AML. As might be expected, class predictors employing 10 to 50 genes were not highly accurate in cross-validation. For example, a 10-gene predictor yielded strong predictions (PS>0.3) for only 40% of the samples, and of those, 67% of the predictions were incorrect. Similarly, a 50-gene predictor yielded strong predictions for 27% of the samples, and 75% of these predictions were incorrect.